CEIC

The California Environmental Information Catalog

http://ceic.resources.ca.gov

Roger Kunkel
California Resources Agency
CERES Program

CEIC is an On-line Metadata Catalog. What is Metadata?

- **Metadata is information that describes a resource:**
  - **What type of resource (e.g. digital image, hard copy report, etc.)?**
  - **Who created it and for what purpose?**
  - **When was it created?**
  - **Where (what is the geographic context)?**
  - **How to obtain it (e.g. file download, mail hard copy)?**
  - **Much more.**

Functions of Metadata

- **Discover Resources**
- Manage Documents
- Control IP Rights
- Identify Versions
- Certify Authenticity
- Indicate Status
- Mark Content Structure
- **Situate Geospatially**
- Describe Process

Standards

- FGDC - Content Standard for Digital Geospatial Metadata
  - The Content Standard for Digital Geospatial Metadata (CSDGM), Vers. 2 (FGDC-STD-001-1998) is the US Federal Metadata standard. The Federal Geographic Data Committee originally adopted the CSDGM in 1994 and revised it in 1998. According to Executive Order 12096 all Federal agencies are ordered to use this standard to document geospatial data created as of January, 1995. The standard is often referred to as the FGDC Metadata Standard and has been implemented beyond the federal level with State and local governments adopting the metadata standard as well.

- US Profile of the (ISO) international metadata standard, ISO 19115
  - The international community, through the International Organisation of Standards (ISO), has developed and approved an international metadata standard, ISO 19115. As a member of ISO, the US required to revise the CSDGM in accord with ISO 19115. Each nation can craft their own profile of ISO 19115 with the requirement that it include the 13 core element. The FGDC is currently leading the development of a US Profile of the (ISO) international metadata standard, ISO 19115.

- Dublin Core
  - Simple DC – 15 elements for general anotation of a resource.

These are typically expressed in XML for machine to machine communication.

Cross-walking Standards

Controlled Vocabularies
Improve quality of search results

- Keyword Thesauri
  - CERES, National(UK) Institute for Health and Clinical Excellence (NICE)
- Place Names
  - GNIS, Counties, etc.
- Resource Types
  - CERES, ISO, etc.

Available Tools

- Free/Open Source
    - Tkme, MP
    - MetaScribe
    - MERMAid
    - Metavist 2005
    - CEIC
        - On-line metadata management
    - CalEDLN
        - Under development

- Commercial
    - ArcCatalog
    - SMMS

The Publish and Harvest Model

- Stewards of resources create and publish their own metadata.
- Aggregators harvest records and provide search and discovery tools.
- Web Analogy –
  - Highly successful publishing model – with low participation cost.
  - HTML / Commercial Search Engines
    - Not enough structure = noise / false positives
  - XML /   Focused Indexers
    - Thoroughly structured = higher quality results

What is CEIC?

- **A web based catalog of data resources for California.**

- **CEIC stores "metadata" not the data itself.**

- **CEIC uses a metadata standard called:**

Content Standard for Digital Geospatial Metadata (**FGDC-STD-001 June 1998**)
Commonly called the "FGDC metadata standard"*

*Note that other standards exist, e.g. ISO 19115

Who Uses CEIC?

- **The Catalog has been developed through a collaborative effort with the <u>California Geographic Information Association</u>, <u>California Environmental Resources Evaluation System</u>, and the <u>Federal Geographic Data Committee</u>.**

- **Participants include cities, counties, utilities, state and federal agencies, private businesses and academic institutions that have spatial and other types of data resources.**

- **Currently 250 organizations have created catalogs with a total 5400 datasets.**

CEIC and Other Tools

- CEIC interoperates with other metadata tools:
  - ESRI ArcCatalog
  - Intergraph SMMS
  - Others using the FGDC standard and importing/ exporting in XML format.

CEIC

xml

xml

Other Systems
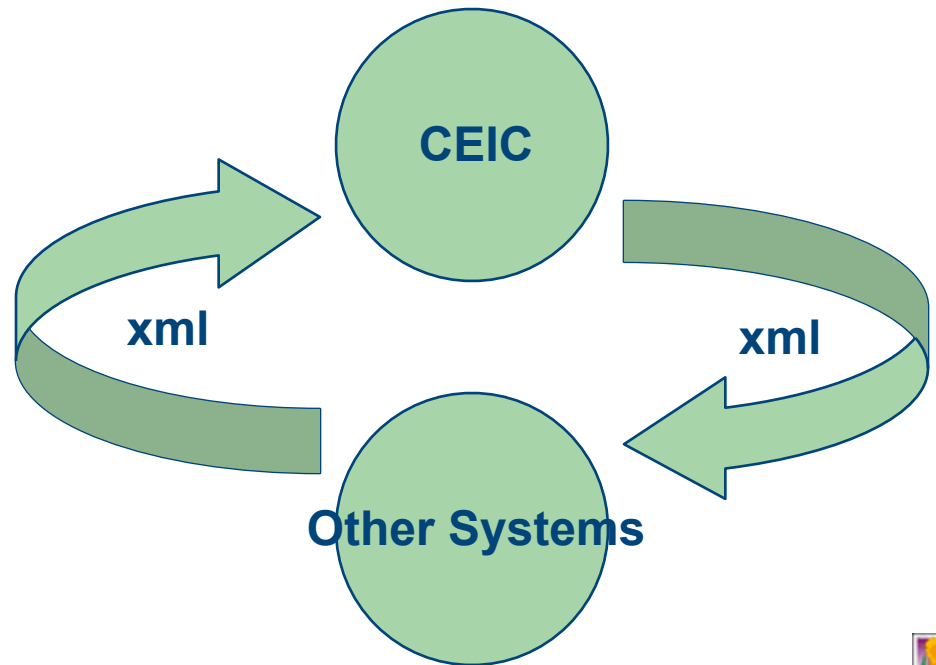
Contributing to CEIC

- On-line System
  - Anyone can create a new catalog
  - Fill out on-line forms of min. FGDC or upload conforming XML files
  - System is moderated for appropriate content
- Publish and Harvest Network
  - Publish conforming XML records to a web accessible directory
 - Register as a CEIC node

General CEIC Architecture
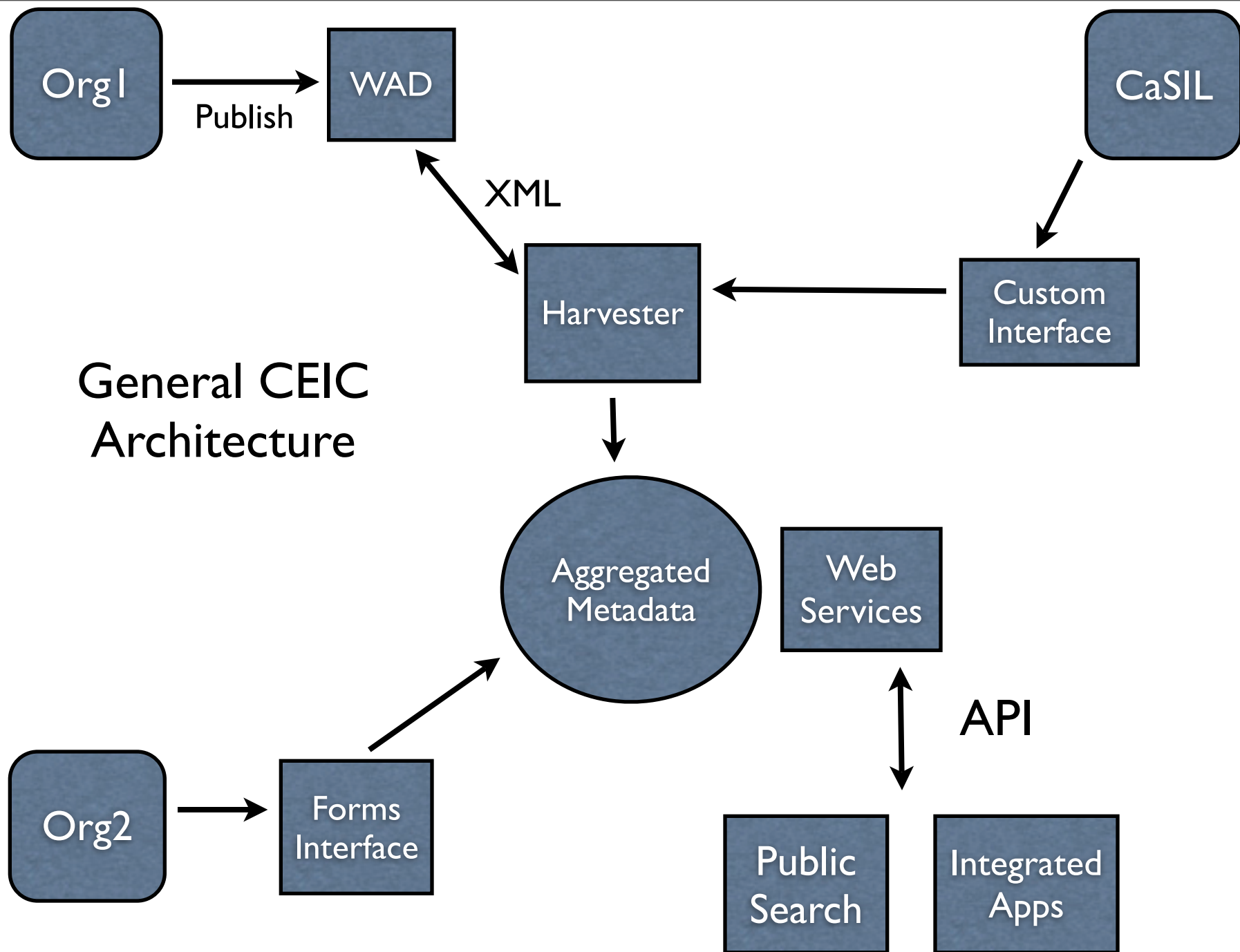
CERES Open API

- Restful Http JSON Format
- Allows application level integration with CEIC
  - Query with Term, Type, and Spatial Parameters
  - Vocabulary for Organization, Theme, Place
  - Spatial Intersections and Relations
- Planned Functions
  - Authentication
  - Anti Duplication Algorithm
  - Realtime XML Metadata Creation & Editing

FGDC Support in CEIC
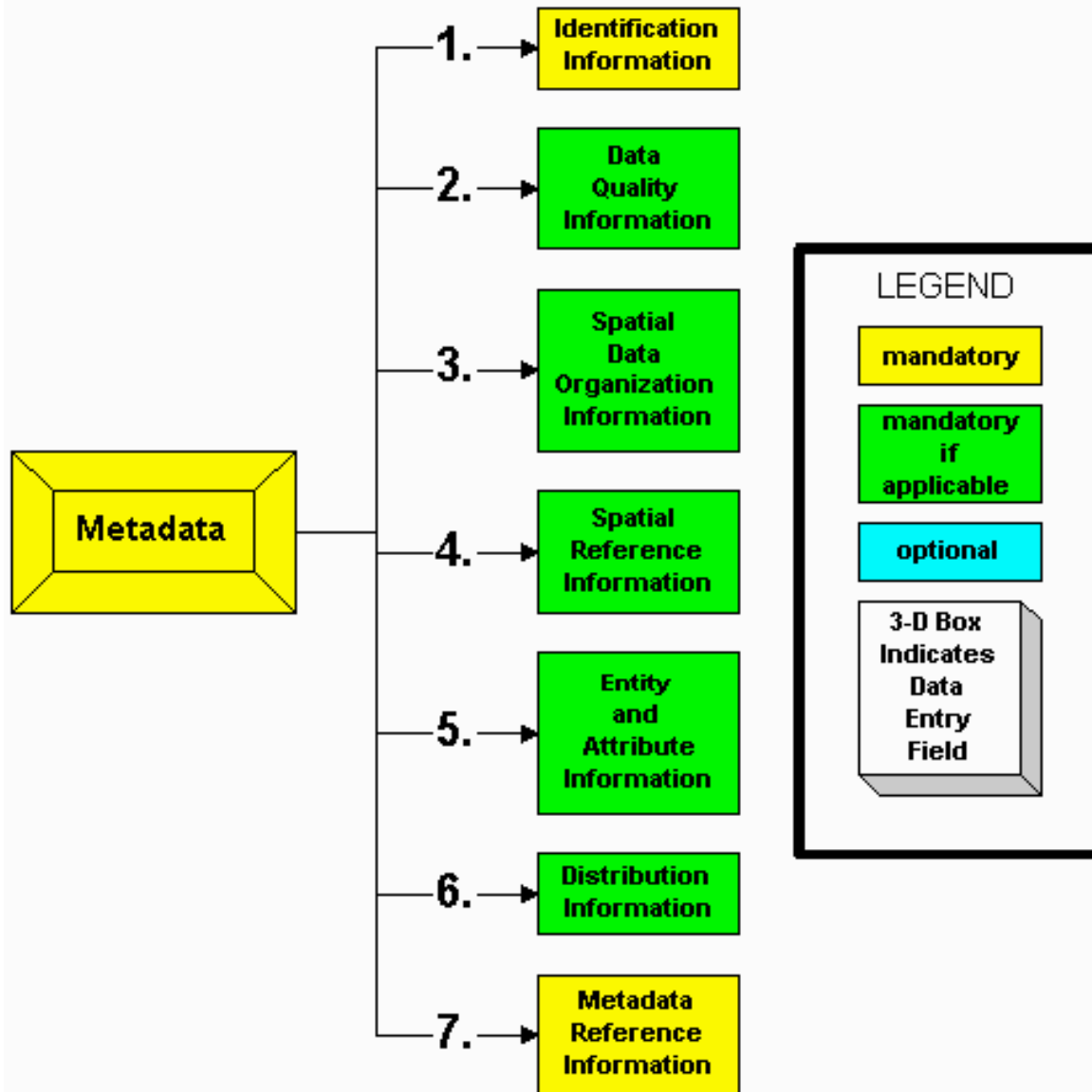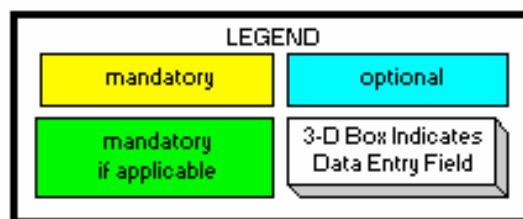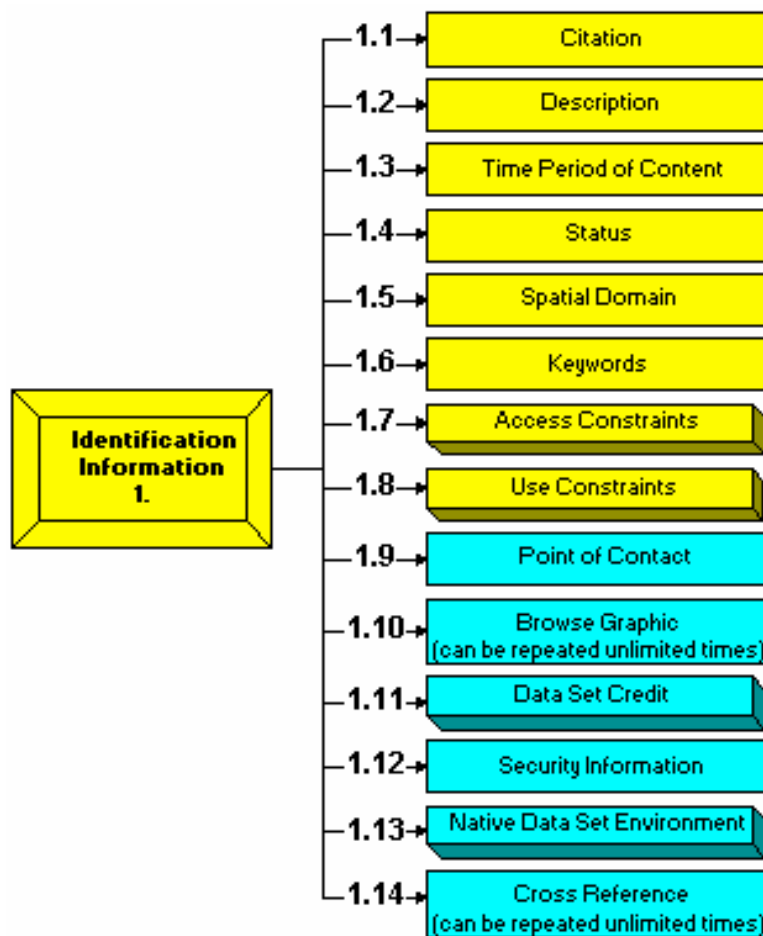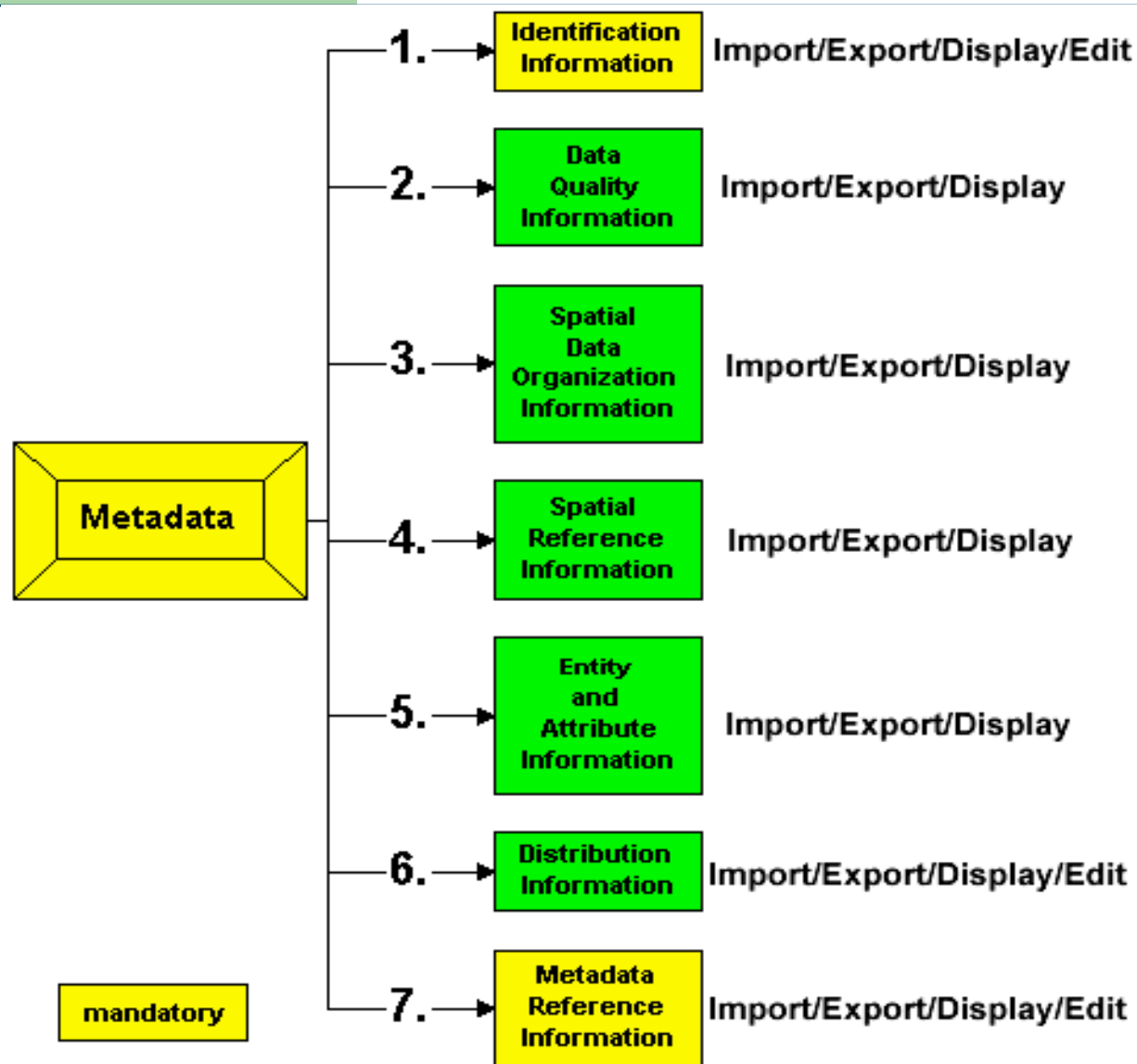
Adding Datasets

Option 1: To add a new dataset using Resources Agency Data Product form, choose "Continue".

[Continue]

Option 2: To add a new dataset using the Standard FGDC Metadata entry form, choose "Continue".

[Continue]

Option 3: To upload a FGDC compliant XML dataset, specify your local file and choose "Submit".

[   ] [Browse...]

[Submit]

[Cancel]

Starting a Catalog

CEIC

In order to start a new catalog you will need to provide a point of contact and some basic information about your catalog.

After submitting this form you can proceed to add datasets to your catalog. Until your catalog and contact information has been reviewed and verified logged in using your provided password.

An email notification will be sent when your catalog has been reviewed and you will then have the option of making your catalog publicly viewable.

Contact Name (*):

Contact Email (*):

Catalog Name (*):

Organization (*):

Organization URL: http://

Metadata Harvest URL: http://

Password (min 8 chars) (*):

Retype Password (*):

A brief description of your catalog:

Submit   Cancel

Finding Datasets

CEIC

The California Environmental Information Clearinghouse (CEIC) is an information clearinghouse developed by the California Resources Agency established to facilitate access to a variety of electronic data describing California's rich and diverse environments. (more...)

Search Term:

stream   (search) (clear)

☐ Exact Phrase   Advanced Search

Information Types:
☑ 🖼 GIS Data   ☐ ⚙ Projects   ☑ ❓ Other

Subjects:
☑ Biota/Environment   ☑ Agriculture
☑ Ocean/Water   ☑ Society/Infrastructure

Location Search: ☐ All ⦿ Map Area Only

[Map] [Sat] [Hyb]

Zoom & drag map to change location

Matching Cataloged Resources *(mouse over for location centerpoint - click for entry detail)*

- 205J Watershed Analysis Project (document)
- Friesen Lakes- drainage (vector digital data)
- Chinook, Critical Habitat, NOAA - Central Valley [ds125] (database)
- **Stream** Habitat Reach Summary - Russian River [ds77] (database)
- Steelhead, Critical Habitat, NOAA - Central Valley [ds123] (database)
- Steelhead - Point Survey [ds184] (database)
- Constructed Fish Passage Barriers (vector digital data)

Frequently Accessed

CEIC

Frequently Accessed Data (FAD)

Most data are maintained in an Albers projection to enable overlay and integration functions, and are organized by county, quad area, or statewide units.

Administrative Divisions

ARB California AirBasins

ARB California AirDistricts

California Park Boundaries 2007

Congressional District Regions 2001

Congressional District Regions 2001 with Additional Tiger Information

County Boundaries (1:24000)

Federal Water Districts - Mid-Pacific Region

Private Water Districts for California 1:24,000-scale

State Water Districts for California

Cultural Geography

California General Plans

California General Plans with Rural Residential Areas

Census 2000 Block Data

Census 2000 Block Group Data

Census 2000 Place Names